Experimental games for the design of reputation management systems

by C. Keser

Trust between people engaging in economic transactions affects the economic growth of their community. Reputation management systems, such as the Feedback Forum of eBay Inc., can increase the trust level of the participants. We show in this paper that experimental economics can be used in a controlled laboratory environment to measure trust and trust enhancement. Specifically, we present an experimental study that quantifies the increase in trust produced by two versions of a reputation management system. We also discuss some emerging issues in the design of reputation management systems.

Trust is at the root of almost any economic or personal interaction. We need to trust the government, the credit card company, and the car dealer. Fukuyama argues that trust, loosely defined as the expectation of trustworthy and cooperative behavior of others, impacts the performance of all social institutions, including firms, and thus the overall economic performance of a country.[1] Individuals in higher-trust societies spend less to protect themselves from being exploited in economic transactions. Trust is an economical substitute for extensive contracts, litigation, and monitoring in transactions and thus economizes on transaction costs. Knack and Keefer provide empirical evidence for Fukuyama's hypothesis, showing that differences in trust across countries help explain differences in investment and economic growth.[2] They measure trust in a country based on an indicator from the World Values Surveys[3] conducted in 1981 and 1990–1991. The indicator is the percentage of respondents from that country who answered positively when asked "*Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people*?"

Based on the same trust indicator, Huang et al. show that more trusting countries tend to exhibit higher levels of Internet penetration.[4] Litan and Rivlin[5] and Varian et al.[6] predict that, due to reduced transaction costs associated with production and distribution of goods and services, the Internet will positively affect productivity in the coming years. Because productivity gains are higher when the level of Internet adoption is higher, low trust countries, most of which tend to be of low and middle income, are doubly penalized in terms of economic growth. First, they are penalized for low trust in terms of investment and growth impact, and then again through lower adoption of growth-enhancing technologies. In other words, we observe a developmental-cum-digital divide illustrated in Figure 1. Although this is bad news for low-trust countries, we show in this paper that there are ways to alleviate the problem.

In the next section we define trust and describe a number of factors that can affect it. Then, we describe a case of successful trust enhancement by reputation management in the context of Internet transactions—the Feedback Forum[7] of eBay Inc. We point out some shortcomings of the Feedback Fo-

©Copyright 2003 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

 0018-8670/03/$5.00 © 2003 IBM

rum and the difficulties involved in experimenting with the operational system. In the following section we present a new methodology that can be used in an experimental economics laboratory to address such questions through experiments with reputation management systems. We also present results that quantify the increase in trust caused by two versions of a reputation management system. In the last section we summarize our results and raise a number of questions for further study.

Trust enhancement and the eBay example

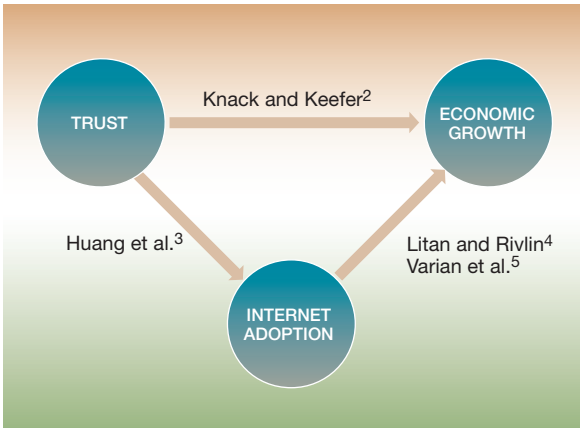
Trust often helps solve problems caused by social uncertainty, as when we are not able to determine the intentions of people or organizations that have incentives to act against our best interest. Following Yamagishi and Yamagishi, we define trust as the expectation of other persons' goodwill and benign intent, implying that in certain situations those persons will place the interests of others before their own.[8]

Yamagishi and Yamagishi distinguish trust from assurance, which they define as the expectation of another person's goodwill and benign intent based on the knowledge of the incentive structure surrounding the relationship. They give the example of a Mafia family member who can expect his trading partners will not cheat, not because they are benevolent people but because they are aware of the consequences. In other words, the partners behave trustworthily because it is in their own interest. Assurance can thus complement or substitute for trust.

A related trust substitute is commitment. Maintaining long-term relationships with loyal partners rather than making deals with new partners is a kind of commitment where incentives for non-cooperative behavior are reduced.[9] In such relationships it is mutual assurance based on the nature of the relationship rather than trust that leads to cooperative behavior. This was demonstrated, for example, by Axelrod,[10] by Selten, Mitzkewitz and Uhlich,[11] and by Keser[12]; they examined human strategies in repeated "social dilemma situations" where the individual payoff-maximizing non-cooperative behavior leads to socially inefficient outcomes. They observed that people often actively attempt to establish and maintain mutual cooperation when they expect to repeatedly interact with each other.[13]

Thus, in early interactions the participants signal their willingness to cooperate and then use reciprocity—cooperate if the others cooperate and avoid cooperation if the others failed to cooperate in the previous interaction—as an instrument to induce cooperation. Following such a strategy typically pays for an individual involved in repeated encounters with others. Keser and van Winden show that in a social dilemma situation people who repeatedly interact with the same people cooperate significantly more than those whose partners change randomly.[14]

Figure 1 The development-cum-digital divide



Familiarity also can be seen as a complement to trust.[15] Familiarity deals with understanding the current action of another person, whereas trust deals with the belief in a future action by that person. Indeed, the latter may often be based on familiarity. Familiarity with Amazon.com**, for example, means knowing how to search for books and information about them, and knowing how to order these books through the Web interface. Trust in Amazon.com, Inc. might entail willingness to provide credit card information based on the belief that the information will be properly used. Though trust and familiarity are distinctly different, they are related.

Another trust complement, or trust-enhancing factor, is reputation. Reputation plays two different roles in social interactions involving trust. The first role is informational. It makes a person trust more when given favorable information about the business partner. Trust has been defined above as the expectation that others will show goodwill in their dealings with us. Lacking perfect information about others' intentions, we thus evaluate their intentions from the available information, such as their reputation.

The second role reputation plays is as a tool for disciplining or restraining, in order to control dishonest behavior. This aspect of reputation makes the targeted party act in a more trustworthy way. In other words, a reputation management system may enhance trust through the creation of assurance.

In e-business we observe successful trust enhancement by reputation management systems. The most popular of those is eBay's Feedback Forum.[16] At eBay, anonymous individuals spread over the globe may buy and sell almost anything, from PEZ** dispensers to Ferraris and castles. With nearly 50 million registered users, 170 million transactions, and $9.3 billion worth of goods sold in 2001, it is the largest of the informal on-line markets.[17] These numbers are impressive, given the risks involved in trading in such a market. Typically there is no opportunity for the buyer to inspect the prepaid item before delivery, and if the quality is unsatisfactory, it may be impossible to track down the seller. Even worse, the buyer has no guarantee that the item will be delivered at all. On the other hand, if the seller chooses to deliver before receiving the payment, there are similar risks involved. To put it differently, both parties involved in a trade might be tempted to cheat. eBay has a fraud protection program that covers losses for up to $200. Beyond that, however, if users do not make use of costly escrow services offered by eBay, they must have a high level of trust when they engage in transactions in this informal on-line market.

To enhance the trust in, and trustworthiness of, its users, eBay created the Feedback Forum. The participants in a transaction are asked to rate each other by submitting a comment and a rating. The rating takes one of three values: "+1" for a positive comment, "−1" for a negative comment, and "0" for a neutral comment. All ratings that an eBay user receives from distinct other users are summed up into a Feedback Rating number. This number is attached to each participant, be it a seller or a bidder. A user who accumulates 195 positive comments and no negative comments has a Feedback Rating of 195. However, a user with 223 positive and 28 negative comments has the same Feedback Rating. A user whose Feedback Rating number drops to −4 is suspended from further participation (recently eBay modified the Feedback Rating to include the percentage of positive comments). The Feedback Rating is part of the user's Feedback Profile, which can be obtained by clicking on the user's Feedback Rating. A user's Feedback Profile includes all comments for that user, the distribution of all previous ratings received from distinct other users, as well as the distribution of recently received ratings over the past seven days, past month, and past six months.[18]

Recently, a number of empirical studies addressed the question as to whether the prices that sellers obtain on eBay are correlated with their reputation. Kalyanam and McIntyre examined auctions of Palm Pilot personal digital assistants,[19] Houser and Wooders examined auctions of Intel Pentium** III processors,[20] and Lucking-Reiley et al. examined collectible coin auctions.[21] All these studies come to the conclusion that buyers are willing to pay more for goods purchased from a seller with a good reputation. Resnick et al. conducted a field experiment in which they sold matched pairs of items (batches of vintage postcards), one half through a seller with a very high reputation and the other half through a newcomer.[22] Then, they compared sales under newcomer identities with and without negative feedback. They observed that the established seller fared better than the newcomer. Furthermore, for newcomers, one or two negative comments did not affect the price. These empirical and experimental field studies thus show that the reputation heterogeneity created by eBay's reputation management system leads to more price dispersion than one would expect when one considers the low search costs to buyers.[23] We are aware of no field study, however, that examines the impact of specific aspects of reputation management systems on their effectiveness, and in particular on trust and trustworthiness.

Although the reported fraud rate at eBay is as low as one percent of all transactions, there are recurring incidents of fraud that may be due to shortcomings in the reputation management system. Dingledine, Freedman and Molnar, for example, describe incidents on eBay in which sellers had built a reputation through a large number of low-value transactions. They then proceeded to offer a number of high-value items, receive payment for these items, and then disappear.[24]

Such incidents indicate that there are drawbacks in the functioning of the current Feedback Forum. The Feedback Rating sums up a user's ratings received both as seller and as buyer. It appears, however, that it is easier for a buyer than for a seller to receive positive evaluations. eBay might consider breaking down the single Feedback Rating number into a set of numbers. The breakdown could be not only according to the role in the transaction (buyer or seller)

but also by product category. A seller of PEZ dispensers may enjoy an excellent reputation among PEZ dispenser buyers and, at the same time, might not be well thought of by car buyers. Building a reputation on low-value deals only to misuse it on high-value deals could also be discouraged by modifying the reputation function: high-value deals, for example, could be weighed more heavily than low-value deals in the calculation of the Feedback Rating. Fraud can also be discouraged by making it more difficult for traders to change their identity. A user with a negative reputation has, in the current version of the Feedback Forum, an incentive to show up as a newcomer with a new identity and the neutral reputation that comes along with it. Without going into further detail, we conclude that many potentially beneficial modifications could be made to eBay's Feedback Forum. One cannot always foresee the implications such modifications have on the behavior of users and the general performance of the reputation management system. Experimenting in the field can be costly and time-consuming. Furthermore, the operational environment limits the extent to which we can modify it. We thus propose to examine the effectiveness of different rating systems in controlled laboratory experiments. In the following section, we present a way to measure trust, trustworthiness, and the enhancement of both through reputation management in an experimental economics laboratory.

Designing reputation mechanisms

In a simulated e-business environment, we consider the interaction of individuals in a situation involving trust (of buyers) and trustworthiness (of sellers). We do not attempt to reproduce the eBay environment, as it involves extraneous issues that, at this point, we do not need.[25,26,27] In the experimental economics literature, Berg, Dickhaut, and McCabe introduced the following "investment game," often called the "trust game."[28]

The two players in this game—let us denote them as buyer and seller—are endowed with ten dollars each. Buyer and seller may interact according to the following rules, of which they both have full knowledge. The buyer may send (invest) part or all of his endowment to the seller, but he need not send anything (we use "he" to refer to either "he" or "she"). The amount invested by the buyer will be tripled (additional funds are injected into the transaction), so that the seller will receive three times the amount invested by the buyer.[29] The seller then has the opportunity to return part or all of the amount received, but need not return anything. Then the game is over.

Suppose that the two players in this game, as typically assumed in economic theory, are driven by self-interest and are striving to maximize their personal payoffs. Suppose further that each assumes the other's objective is the same—maximizing the personal payoff. The game can easily be solved by backward induction: a seller striving to maximize personal payoff will not return anything to the buyer, whatever amount the other invested. The buyer, also striving to maximize personal payoff and anticipating zero return from the seller, will not send anything to the seller. Thus, economic theory predicts zero flow of money in this game.

We propose a way to design reputation management systems by measuring trust and trustworthiness in an experimental economics laboratory.

In the experimental laboratory, as in real life, we often observe people behaving less selfishly than postulated by economic theory. Berg, Dickhaut, and McCabe[28] conducted laboratory experiments on the trust game under strong anonymity conditions, ensuring that neither the participants nor the experiment monitors can associate an action with a specific individual. They gathered a number of participants, half of them as buyers and the other half as sellers. Buyers and sellers were placed in different rooms and never saw each other. To conduct the experiment, an envelope-mailbox system was used. Each buyer had the opportunity to place any amount between zero and ten dollars in an envelope, which he then deposited in a mailbox that he shared with an unidentified seller. The experiment monitor added funds to triple the amount in each of the envelopes before the sellers accessed their mailboxes to pick up the envelopes. Each seller then had the opportunity to return any amount between zero and the received amount; this was placed in the envelope and returned to the mailbox to be picked up by the buyer with whom he shared the mailbox. Participants in these experiments were undergraduate students from the University of Minnesota playing

this game for real money. More than 90 percent of buyers sent money to sellers ($5.16 on average) and about 45 percent of sellers who received some money from the buyer returned a fraction of it ($4.66 on average). The amount invested by the buyer provides a measure for the buyer's trust in the seller's goodwill, whereas the fraction returned by the seller measures the seller's trustworthiness. Thus, the experiments provide evidence that people, to some extent, do trust in the goodwill of others. Furthermore, to some extent people are trustworthy: the sellers returned, in aggregate, almost as much as the buyers invested. Nevertheless, in this game trust does not really pay for the buyer. The sellers tend to keep the entire surplus created by the buyers' investments for themselves, which implies an inequitable payoff distribution between buyers and sellers. Note also that the observed level of trust is still far below the maximum level of efficiency: the sum of payoffs to the two players would be highest if the buyer invested the entire endowment. In other words, only full trust could lead to the maximally efficient outcome. In the case of full trust, any return by the seller greater or equal to ten dollars implies a Pareto efficient situation defined as one in which none of the players could be made better off without reducing the payoff to the other player.

The results of Berg, Dickhaut, and McCabe[28] have been replicated in a large number of experimental studies conducted in various countries. The results are qualitatively robust even though the level of trust varies across countries. Interestingly, the observed differences between countries are in line with the trust indicator in the World Values Survey.[3] Willinger et al., for example, find that Germans trust significantly more than the French.[30] Buchan, Croson, and Dawes observe a higher trust level among Americans and Chinese than among Koreans and Japanese.[31] Ensminger reports on a study involving norms of altruism, trust, and cooperation of people in New Guinea, the Amazon rain forest, Kenya, and in urban and rural Missouri.[32] The results indicate that in the smaller-scale societies of the developing world the level of trust is lower than in the U.S.

The auctioning aspect having been removed, the situation set up by the trust game is similar to trading in on-line markets. The focus is on issues of trust and trustworthiness. If I buy PEZ dispensers at eBay, I face risks similar to those of the buyer in the trust game: Is the quality as described? Will the packaging be appropriate? Will the delivery be timely? Will the item be delivered at all? The seller on eBay may be viewed as the seller in the trust game as far as trustworthiness is concerned. Although the trust game, due to the tripling of the buyer's investment, does not directly correspond to the eBay transaction, it captures in a simple way its major issues of trust and trustworthiness. The tripling of the investment corresponds to gains from trade in a market context. As discussed in the previous section, it is to be assumed that at eBay the buyer's risks with respect to the seller's trustworthiness are somewhat mitigated by the Feedback Forum.[33] On one hand, the seller's reputation is likely to influence the buyer's expectation of trustworthiness; on the other hand, the fear of a negative rating makes sellers behave more trustworthily.

We describe here several computerized laboratory experiments that examine the functioning of reputation management systems (this paper is based on Reference 34 which includes additional details about our experiments). The big advantage of using the laboratory experimental method, compared to field studies or experiments, is that we can directly compare a situation without a reputation management system to situations with specific reputation management systems. We can thus measure to what extent each reputation management system impacts trust and trustworthiness. We have conducted three different experiments. All of the experiments are based on a twenty-fold repetition (over intervals known as "periods") of the trust game. At the beginning of each period, each player receives 10 Experimental Currency Units (ECUs). ECUs are converted to Canadian dollars at the end of the experiment at the rate of Can$0.07 per ECU (all amounts below are ECUs, unless marked otherwise).

The participants remain in the same role of either buyer or seller over all 20 periods. The pairing of a buyer with a seller is random in each period, but the same two players never meet in two consecutive periods. The first experiment, called the *baseline experiment*, is based on the trust game as previously described. The other two experiments are based on a modified trust game involving a reputation management system as described below.

At the end of each period the buyer, after receiving the amount returned by the seller, is asked to rate the seller's cooperation as *positive*, *neutral* or *negative*. The rating is disclosed to the seller. From the second period on, at the beginning of each period, the buyer is given information about the seller's past ratings, as follows. In the second experiment, named

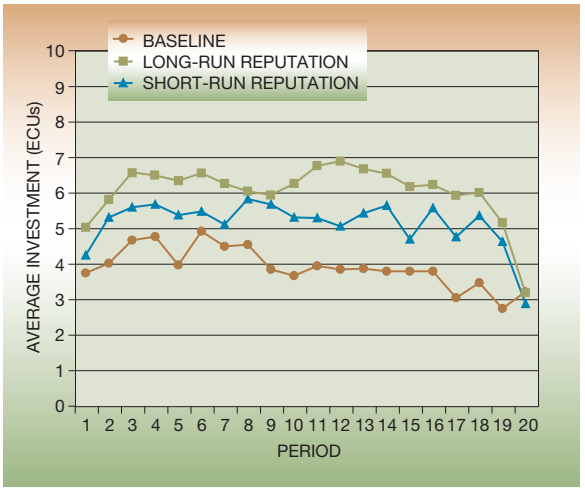
the *short-run reputation experiment*, the buyer is given the seller's most recent rating. In the third experiment, named the *long-run reputation experiment*, the buyer is given the most recent rating and also the distribution of all previous ratings. Note that this information is similar to a user's Feedback Profile on eBay.

The experiments involved 320 student volunteers from various departments at a number of universities in Montreal. Thirty-two experimental sessions were held, eight for the baseline experiment, and 12 for each reputation experiment. Each session had 10 participants—five buyers and five sellers. Buyer-seller pairs were randomly selected in each of the 20 periods, under the constraint that no two participants should be paired in two consecutive periods. Note that the random matching implies that the aggregate behavior in a session represents an independent observation on which non-parametric statistics will be based.[35,36] The rules of the game were identical in each of the 20 periods. The payoffs to a participant over the 20 periods were added up to determine the total individual earnings in the experiment. An experimental session lasted about 1.5 hours and the average earnings were Can$30 (including a show-up fee of Can$5). The experiments were conducted in French in the computerized experimental economics laboratory at CIRANO[37] in Montreal. Instructions were distributed and then read aloud. To qualify, all participants had to correctly answer a number of questions testing their understanding of the instructions. Payment was made in private at the end of the experiment.

The results of the baseline experiment are in line with previous experimental results. In our baseline experiment the buyers' average investment (our measure of trust) is 3.91, and the sellers' average return is 3.3, which amounts to a *relative return* with respect to the received amount (our measure of trustworthiness) of 33 percent. In other words, in aggregate, the buyers receive a little less than the invested amount back, which implies that the sellers keep not only the entire surplus but also a small part of the buyers' investment for themselves. While trustworthiness is relatively stable over time, we observe a statistically significant decrease in the trust level from the first 10 periods to the last 10 periods of the game (two-sided Wilcoxon signed ranks test, 5 percent level).[35,36]

The experiments show the use of reputation management systems has significant effects on trust and trustworthiness. When compared to the baseline results, the short-run reputation management system increases trust and trustworthiness by more than 30 percent (two-sided Mann-Whitney U-tests, 10 percent and 5 percent level, respectively): it increases trust to 5.15 and trustworthiness to 46 percent. The long-run reputation management system is even more efficient, increasing both trust and trustworthiness by more than 50 percent (two-sided Mann-Whitney U-tests, 1 percent level); it increases trust to 6.05 and trustworthiness to 49 percent.

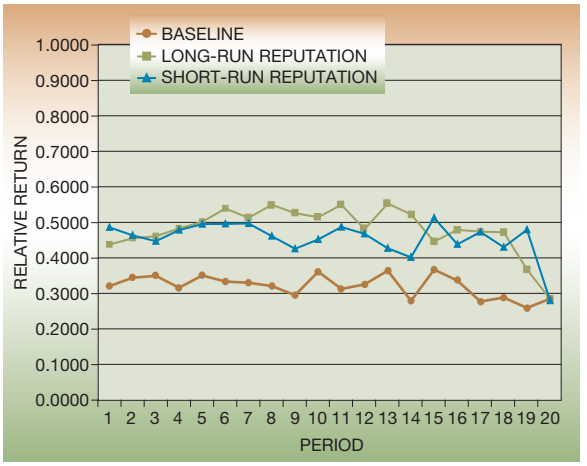
Figure 2 Buyers' investment (trust measure) over time



The variation of trust over time in the three experiments is shown in Figure 2, whereas Figure 3 shows the variation over time for trustworthiness. As these figures illustrate, in aggregate, with a reputation management system in place, both trust and trustworthiness are always higher than in the baseline experiment, except for the final period. The dramatic drop in trust and trustworthiness toward the end of the interaction is a typical end-game effect, as observed in the vast experimental literature on the prisoner's-dilemma type of game with finite repetitions, where cooperation tends to break down toward the end.[38]

In terms of payoff, the major winners from the use of a reputation management system are the buyers. Indeed, with a reputation management system in place, sellers are forced to return a larger fraction of the buyer's investment in order to get a positive rating. The buyers' average (per period) payoff increases from 9.80 in the baseline to 11.95 in the short-

Figure 3 Sellers' relative return (trustworthiness measure) over time



run and to 12.83 in the long-run reputation experiment. Both of these increases are significant at the 1 percent level (two-sided Mann-Whitney U-tests). Sellers also earn higher payoffs than in the baseline experiment: the sellers' average (per period) payoff increases from 17.93 in the baseline to 18.30 in the short-run and to 19.27 in the long-run reputation experiment. Nevertheless, these increases are not statistically significant if a 10 percent significance level is required (two-sided Mann-Whitney U-test).

Sellers care very much for their long-run reputation when it is at stake because buyers' trust increases significantly with the seller's long-run reputation. Assigning a value of +1 to each positive rating, a value of 0 to each neutral rating, and a value of −1 to each negative rating, we define a seller's long-run reputation as the sum of this seller's previous ratings. Based on this metric, we observe that a buyer's investment tends to be higher if the seller's long-run reputation is positive rather than neutral, or neutral rather than negative (two-sided Wilcoxon signed ranks tests, 1 percent level). Short-run reputation also matters, but when information on the long-run reputation is available, additional short-run reputation affects the buyers' trust significantly only when it is positive (two-sided Wilcoxon signed ranks test, 1 percent level). When short-run reputation is the only information available, a positive reputation significantly increases the buyers' trust while a negative reputation significantly decreases it (two-sided Wilcoxon signed ranks tests, 1 percent level in the comparison of positive and neutral reputation and 5 percent level in the comparison of neutral and negative reputation).

To summarize the results, the introduction of a reputation management system significantly increases both the transaction volume (trust) and the level of trustworthiness. The maximum payoff for a buyer-seller pair, taken as the sum of the two payoffs, occurs when the buyer invests the entire allowance of 10 ECU, and results in a total payoff of 40 ECU for the pair. Defining efficiency as the total payoff for the pair as a percent of the maximum payoff, we observe that the average efficiency increases from 69 percent in the baseline to 76 percent in the short-run and to 80 percent in the long-run reputation experiment.[34] The introduction of a reputation management system leads to a Pareto improvement as both player types earn higher payoffs. At the same time, the payoff distribution between buyers and sellers becomes more equitable because in this environment engaging in a transaction that requires trust does pay (in aggregate). This is not the case in an environment without a reputation management system. Comparing the two reputation experiments, we observe that the use of the long-run reputation leads, in the intermediate phase of the interaction, to more trust and trustworthiness than the short-run reputation.

Conclusion

We started our discussion with the observation that trust plays an important role at many levels: for individual transactions, for businesses and organizations at the national level, and for the global economy. Trust varies across countries, but there are ways for low-trust countries to substitute for or to enhance trust. In e-business we observe successful trust enhancement through the use of reputation management systems. Nonetheless, currently there are neither widely used reputation management systems (each on-line market uses its own design) nor any universally accepted guidelines for designing effective reputation management systems.[39]

We have proposed the use of experimental game theory in the design of efficient reputation management systems. The experiments presented in this paper examine the effectiveness of two reputation management systems, of which one, the long-run reputation management system, is very similar to eBay's Feedback Forum. The results suggest that if eBay had not introduced their reputation management system,

they would have experienced less growth and more fraud. The use of the long-run reputation model appears preferable to the short-run reputation model. It remains an open question as to whether eBay could do even better at enhancing trust and trustworthiness by using an enhanced reputation management system. Should they use, for example, a system in which each rating that a user receives is weighted by the value of the transaction? Should the raters' reputation be used to weight the ratings? How important is it that eBay users have no doubts about the technical reliability of the Feedback Forum? How do communities of interest on eBay influence trust and reputation formation? Laboratory experiments provide us with a controllable, fast, and relatively inexpensive means of addressing all these questions. As such, the laboratory is an important complement to theoretical analyses and field studies.

Acknowledgments

I thank CIRANO of Montreal for their hospitality and, in particular, Claude Montmarquette and Jacinthe Majeau for their help in running the experiments. Thanks are also due to Robert Baseman, Bill Grey, Jonathan Leland, Mark Nixon, and Carl Singer for their input and many discussions. I am grateful to the anonymous referees for their valuable comments. Prince Stanley developed the experimental software, which is based on z-Tree by Urs Fischbacher at the Institute for Empirical Research in Economics at the University of Zurich.

**Trademark or registered trademark of Amazon.com, Inc., PEZ Candy, Inc. or Intel Corporation.

Cited references and notes

1. F. Fukuyama, *Trust: The Social Virtues and the Creation of Prosperity*, Free Press, New York (1995).
2. S. Knack and P. Keefer, "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," *Quarterly Journal of Economics* **112**, No. 4, 1251–1288 (1997).
3. R. Inglehart, *World Values Surveys and European Values Surveys, 1981–1984, 1990–1993, and 1995–1997*, Social Sciences Data Collection, University of California San Diego (February 2000), http://ssdc.ucsd.edu/ssdc/icp02790.html.
4. H. Huang, C. Keser, J. Leland, and J. Shachat, "Trust, the Internet, and the Digital Divide," *IBM Systems Journal* **42,** No. 3, 507–518 (2003, this issue).
5. R. E. Litan and A. Rivlin, *Beyond the Dot.Coms: The Economic Promise of the Internet*, Internet Policy Institute, Brookings Institution Press, Washington, DC (2001).
6. H. Varian, R. E. Litan, A. Elder, and J. Sutter, *The Net Impact Study*, http://www.netimpactstudy.com/
7. We have to acknowledge that the relationship between the general level of trust in a country and trust in the specific context of an Internet transaction has not yet been firmly established. Survey studies on privacy, an issue very relevant to online transactions, have shown, for example, that even among relatively high trust countries there may be significant differences in privacy concerns. See for example *IBM Multi-National Consumer Privacy Survey*, IBM Global Services (1999), www.ibm.com/services/files/privacy_survey_oct991.pdf.
8. T. Yamagishi and M. Yamagishi, "Trust and Commitment in the United States and Japan," *Motivation and Emotion* **18**, 129–166 (1994).
9. In Reference 8, Yamagishi and Yamagishi argue that this is typical for Japanese firms.
10. R. Axelrod, *The Evolution of Cooperation*, Basic Books, New York (1984).
11. R. Selten, M. Mitzkewitz, and G. Uhlich, "Duopoly Strategies Programmed by Experienced Players," *Econometrica* **65**, 517–555 (1997).
12. C. Keser, *Strategically Planned Behavior in Public Goods Experiments*, Working Paper, Scientific Series 2000s-35, CIRANO, Montreal, Canada (2000).
13. In Reference 10, Axelrod considers the so-called prisoner's dilemma game; in Reference 11, Selten, Mitzkewitz and Uhlich describe a Cournot oligopoly; and in Reference 12, Keser deals with a public goods game.
14. C. Keser and F. van Winden, "Conditional Cooperation and Voluntary Contributions to Public Goods," *Scandinavian Journal of Economics* **102**, 23–39 (2000).
15. D. Gefen, "E-commerce: The Role of Familiarity and Trust," *Omega*, *International Journal of Management Science* **28**, 725–737 (2000).
16. eBay, Inc., http://ebay.com/.
17. J. Adler, "The eBay Way of Life," Newsweek (June 17, 2002).
18. Amazon.com uses a different feedback system. Any time a buyer at Amazon.com makes a purchase the buyer is encouraged to rate the seller's performance (zero to five stars, five stars being the highest) and leave a short comment. The average rating accompanies the seller's name in product listings.
19. K. Kalyanam and S. McIntyre, *Returns to Reputation in Online Auction Markets*, presentation to IBM Consulting Group (2001).
20. D. Houser and J. Wooders, *Reputation in Auctions: Theory and Evidence from eBay*, University of Arizona (2000, working paper).
21. D. Lucking-Reiley, D. Bryan, N. Prasad, and D. Reeves, *Pennies from eBay: the Determinants of Price in Online Auctions*, University of Arizona (2000, working paper).
22. P. Resnick, R. Zeckhauser, J. Swanson, and K. Lockwood, *The Value of Reputation on eBay: a Controlled Experiment*, University of Michigan (2002, working paper).
23. J. Y. Bakos, "Reducing Buyer Search Costs: Implications for Electronic Marketplaces," *Management Science* **43**, 1676–1692 (1997).
24. R. Dingledine, M. J. Freedman, and D. Molnar, "Accountability," in: *Peer-to-Peer: Harnessing the Benefits of a Disruptive Technology*, A. Oram, Editor, O'Reilley & Associates, Sebastopol, CA (2001).
25. The design of auctions for online markets is a highly interesting topic that has been addressed by, for example, Roth and Ockenfels[26] and Ockenfels and Roth.[27] The role of trust in the auction framework is part of our agenda for future research.
26. A. E. Roth and A. Ockenfels, "Last-Minute Bidding and the Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon Auctions on the Internet," *American Economic Review* **92**, 1093–1103 (2002).
27. A. Ockenfels and A. E. Roth, *Late and Multiple Bidding in*

Second Price Internet Auctions: Theory and Evidence Concerning Different Rules for Ending an Auction, Harvard Business School (2002, working paper).
28. J. Berg, J. Dickhaut, and K. McCabe, "Trust, Reciprocity, and Social History," *Games and Economic Behavior* **10**, 122–142 (1995).
29. The tripling of the amount invested might be considered as gains from trade.
30. M. Willinger, C. Keser, C. Lohmann, and J.-C. Usunier, "A Comparison of Trust and Reciprocity Between France and Germany: Experimental Investigation Based on the Investment Game," *Journal of Economic Psychology* (forthcoming).
31. N. R. Buchan, R. T. A. Croson, and R. M. Dawes, "Swift Neighbors and Persistent Strangers: A Cross-Cultural Investigation of Trust and Reciprocity in Social Exchange," *American Journal of Sociology* **108**, No. 1, 168–206 (2002).
32. J. Ensminger, "From Nomads in Kenya to Small-Town America: Experimental Economics in the Bush," *Engineering and Science* **2**, 6–16 (2002).
33. The Feedback Forum also mitigates the seller's risks with respect to the buyer's payment.
34. C. Keser, *Managing Brands in e-Business: An Experimental Study in Trust and Reputation Management*, Research Report RC22654, Research Division, IBM Corporation (2002).
35. We use *Statistical Package for the Social Science* (*SPSS*), University Information Technology Services, Indiana University, Bloomington, IN; http://www.indiana.edu/~statmath/index.html.
36. See for example S. Siegel and N. J. Castellan, *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill, New York (1988).
37. The Centre for Interuniversity Research and Analysis on Organizations (CIRANO) is a non-profit corporation for research targeted to helping businesses become more competitive. Its members come from industry, academia, and the government. CIRANO is located in Montreal; http://www.cirano.qc.ca/en/bref.php.
38. R. Selten and R. Stoecker, "End Behavior in Sequences of Finite Prisoner's Dilemma Supergames," *Journal of Economic Behavior and Organization* **7**, 47–70 (1986).
39. P. Kollock, "The Production of Trust in Online Markets," in *Advances in Group Processes* (Vol. 16), E. J. Lawler, M. Macy, S. Thyne, and H. A. Walker, Editors, JAI Press, Greenwich, CT (1999).

Accepted for publication March 24, 2003.

Claudia Keser *IBM T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598 (ckeser@us.ibm.com).* Dr. Keser is a research staff member at the IBM T. J. Watson Research Center, Associate Fellow of CIRANO (Centre Interuniversitaire de Recherche en Analyse des Organisations) in Montreal, and Privatdozentin at the Technical University of Karlsruhe. She received her Doctoral degree in Economics in 1992 at the Rheinische Friedrich-Willhelms University of Bonn, working with Nobel Laureate Reinhard Selten on experimental duopolies with demand inertia. Her research in experimental game theory has been focused on issues of incentives, trust, and cooperation.